DOCUMENT RESUME

ED 457 150 SP 040 252

AUTHOR Snow-Renner, Ravay
TITLE Teachers' Perspectives on Standards-Based Education: Initial Findings from a High-Performing, High-Needs School District.
INSTITUTION Mid-Continent Research for Education and Learning, Aurora, CO.
SPONS AGENCY Office of Educational Research and Improvement (ED), Washington, DC.
PUB DATE 2001-09-00
NOTE 42p.
CONTRACT ED-01-CO-0006
AVAILABLE FROM Mid-Continent Research for Education and Learning, 2550 South Parker Road, Suite 500, Aurora, CO 80014-1678. Tel: 303-337-0990; Fax: 303-337-3005; e-mail: info@mcrel.org; Web site: http://www.mcrel.org.
PUB TYPE Reports - Research (143)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Achievement; *Academic Standards; *Educational Change; Elementary Education; Elementary School Students; Elementary School Teachers; *Poverty; Principals; *Scores; *Standardized Tests; State Standards; Student Evaluation; Teacher Attitudes
IDENTIFIERS *Reform Efforts

ABSTRACT

This study examined elementary teachers' perspectives about key aspects of standard-based education reform. Using survey data from teachers in schools of varying poverty and student achievement levels, the study explored beliefs about elements of standards-based reform and perceptions of principals' emphases on instruction and test scores. Teachers believed that using scores to guide teaching helped student learning in mathematics and language arts. However, they did not feel that externally mandated state and district tests were useful in diagnosing student learning needs. Over half of the teachers did not feel that the state accountability test accurately measured student learning. They believed that emphasis on the state test had led to narrower curriculum and less time spent on content areas not directly tested. Teachers in poorer schools were significantly more positive about the use of test results to improve student learning than were teachers in less impoverished schools. They were also more likely to provide additional learning time for non-proficient students in mathematics and language arts. In extremely impoverished schools, teachers reported significantly more learning time for non-proficient students in mathematics. Charts, tables, and data are appended. (Contains 59 references.) (SM)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 457 150

Mid-continent Research for Education and Learning

Teachers' Perspectives on Standards-Based Education:

Initial Findings from a High-Performing, High-Needs School District

SCOPE OF INTEREST NOTICE

The ERIC Facility has assigned this document for processing to:

In our judgment, this document is also of interest to the Clearinghouses noted to the right. Indexing should reflect their special points of view.

Regional Educational Laboratory
Contract #ED-01-CO-0006
Deliverable #2001-05

Office of Educational Research and Improvement
U.S. Department of Education
Washington, D.C. 20208

prepared by
Ravay Snow-Renner, Ph.D., Researcher

Mid-continent Research for Education and Learning
2500 S. Parker Road, Suite 500
Aurora, CO 80014
303-337-0990 (phone)
303-337-3005 (fax)
www.mcrel.org

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Louis Cicchinelli

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

September, 2001

McREL

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

SP040252

© 2001 McREL

To order copies of *Teachers' Perspectives on Standards-Based Education: Initial Findings from a High-Performing, High-Needs School District*, contact McREL:

Mid-continent Research for Education and Learning
2550 South Parker Road, Suite 500
Aurora, CO 80014-1678
tel: 303-337-0990
fax: 303-337-3005
web site: www.mcrel.org
e-mail: info@mcrel.org

The compilation and dissemination of this report is sponsored by the U.S. Department of Education, Office of Educational Research and Improvement, under contract number ED-01-CO-0006. The content of this publication does not necessarily reflect the views of any individual laboratory, the Department of Education, or any other agency of the U.S. Government.

TABLE OF CONTENTS

Acknowledgements ii

Abstract iii

Introduction 1
The Historical Development of Standards Policy 2
The Current Context for Standards in Policy and Practice 5

Method 7
Sample, Participants, and Measure 7
Analysis 10

Results 10
Teacher Perspectives 10
Perspectives about Testing, Test Use, and Learning Opportunities 10
Perspectives about the Curriculum and the State Test 12
Perspectives about Instruction Based on Standards 13
Principals' Emphases 14
Differences by School Conditions 16
Differences in Teacher Perspectives 17
Differences in Principals' Emphases 18

Discussion 19

References 22

Appendix A. Descriptive information about sampled elementary schools A1
Appendix B. Survey items B1
Appendix C. Description of factor analyses examining teachers' perspectives and principals' emphases C1
Appendix D. Bar charts of significantly different mean ratings on different aspects of standards-based education, by school conditions D1

ACKNOWLEDGEMENTS

The author would like to acknowledge the input of several individuals who have contributed to this study. Nancy Sanders assisted with initial problem definition and direction for the study, and provided input on earlier drafts of this document. Helen Apthorp, Zoe Barley, Ken Howe, and representatives from the school district studied provided feedback and direction about the presentation and organization of the findings described here, based on their readings of early drafts. Further thanks are due to Becky Van Buhler, who assisted in data cleanup, and Mya Glenn, Barb Gaddy, and Barb Aiduk, who helped to prepare the final document. Finally, and most importantly, the author would like to thank all of the teachers in the study sites, who gave most generously of their time and expertise to help with this study.

ABSTRACT

This report describes the initial findings of an exploratory study into teachers' perspectives about key aspects of standards-based education reform. It draws on data collected as part of a broader study of policies and practices in high-performing, high-needs elementary schools in one school district in the Midwest. Using survey data collected from 172 teachers in ten elementary schools of varying poverty and student achievement levels, teachers' beliefs about elements of standards-based reform and their perceptions of principals' emphases on instruction and test scores were explored and analyzed using descriptive statistics. In order to examine possible variations by school conditions (poverty level and achievement), Analyses of Variance were conducted.

Teachers generally reported that the use of test results to guide teaching in their schools had helped student learning in mathematics and in language arts. However, they did not feel that externally mandated state and district tests were useful in helping them diagnose student learning needs. Approximately 60% of teachers also disagreed with the idea that the state accountability test was a good measure of students learning in mathematics or language arts standards. And teachers overwhelmingly reported that an emphasis on the state test had led to a narrowed curriculum and less time spent on content areas that were not directly tested. The idea of a common vision that helped to shape the nature of standards-based instruction was not generally present in these schools. Principals were perceived by their teachers to have a strong emphasis both on test results and on instructional quality, with slightly more emphasis on instructional quality.

Teachers in poorer schools were significantly more positive about the use of test results to improve student learning than teachers in the most affluent schools selected (which only had moderate poverty levels). This was the case both in mathematics and in language arts. Teachers in high-performing, extremely-high-poverty and high-performing, high-poverty schools were also significantly more likely to provide additional learning time for nonproficient students in both content areas than teachers in high-performing, moderate poverty schools.

In schools of extreme poverty, teachers in high-performing schools reported additional learning time for nonproficient students in mathematics — significantly more so than did teachers in low-performing schools. Another difference between high- and low-performing schools serving populations in extreme poverty is that teachers in high-performing sites perceive their principals as having a significantly greater emphasis on matters of instruction than do their counterparts in low-performing sites.

INTRODUCTION

Policies about standards, assessments, and accountability dominate current education reform efforts. Such policies form an important aspect of the context in which teachers, students, and schools must function. However, these policies provide sometimes conflicting messages to educators about instruction. For example, the vision of ambitious changes in instruction for all students spurred the development of initial standards policy work, which included content standards, performance standards, delivery standards, and system performance standards (National Council on Education Standards and Testing [NCEST], 1992). This vision was based in content-specific descriptions of teaching and learning envisioned by reformers in mathematics (e.g., National Council of Teachers of Mathematics [NCTM], 1989, 2000) and science (Rutherford & Ahlgren, 1990), which may be characterized as *standards as the basis of instruction*.

Over the years, however, the original comprehensive vision of a standards-based education system as described by NCEST (1992) evolved into a leaner policy model that consisted primarily of standards and assessments, without sustained attention to delivery standards or system performance standards. Increasingly, this model has taken on characteristics that can best be described as *standards as the basis of accountability*. The perspective on standards-based education that interprets standards as accountability is seen as a stark contrast to the idea of standards as having implications for instruction. This perspective pays little or no attention to the nature of classroom instruction (Simon, Passantino, & Foley, 1998) and primary attention to student achievement results on large-scale tests, generally administered to students at different grades on a set schedule.

In part as a result of these types of disparate messages about the purpose of standards policies, local implementation of standards reforms varies widely. At the school and classroom levels, the ways in which standards are implemented are dependent on several factors including the perspectives that teachers and principals take on what standards-based education means. A considerable body of research indicates that pressure to raise student scores on external tests does not necessarily guide teachers toward the types of practices advocated by early proponents of *standards as the basis of instruction* (e.g., Haladyna, Nolan, & Haas, 1991; Frederiksen & Collins, 1989; Firestone, Mayrowetz, & Fairman, 1998). In particular, in low-performing schools (McGill-Franzen & Ward, 1997) and for students who typically do not do well on tests (McNeil, 2000), there is evidence that pressure to raise scores leads teachers to narrow the enacted curriculum and to teach to the test, rather than to adapt instruction to address the set of high learning expectations for all students contained in the standards.

In considering what implementation of standards means, it is necessary to clarify the definition of implementation and to take into account these ideas when considering instructional change. It is also necessary to examine classroom practices and the extent to which local educators perceive standards as having implications for changing that practice. A perspective on *standards as the basis of instruction* addresses broad and ambitious learning goals that apply to every student. High performance on an external test may serve as one indicator that these goals are being met, but it is only an indicator, not the ultimate goal — which is to help all students learn and grow to their full potential. A perspective that interprets standards *as the basis of*

accountability- on the other hand, does not typically interpret standards as having any meaning beyond improved test performance. Indeed, it creates a situation in which the test items, by default, becomes the standards (Thompson, 2001).

This study examines the ways in which school-level educators reconcile these different perspectives on standards. It uses data from a large-scale survey of elementary school teachers in one school district to explore teacher and principal perspectives on various facets of standards-based education, including external tests, implications of standards for instructional practice, and the use of test results. It draws on quantitative data collected as part of a larger comparative case study of selected elementary schools in one district that has demonstrated high capacity for raising students' state test scores, particularly for its high-needs students. The patterns highlighted in this study provide initial information about the nature of teachers' and principals' shared beliefs, attitudes, and knowledge about standards-based education. They also suggest logical directions for further study, both in terms of approaches to classifying perspectives on standards-based education, and in triangulating these data with qualitative data.

THE HISTORICAL DEVELOPMENT OF STANDARDS POLICY

Standards-based education, as it was initially conceptualized in policy and research documents (Conference Report, 1994; McLaughlin & Shepard, 1995) advocated *ambitious learning goals for all students*. As envisioned by NCEST (1992), a fully developed standards-based education system included

- *content standards* that describe the knowledge, skills, and other understandings that schools should teach and students learn in order for students to attain high levels of competency in challenging subject matter;
- *student performance standards* that define various levels of competence in the challenging subject matter set out in the content standards;
- *school delivery standards* [or opportunity-to-learn standards] developed by the states collectively from which each state could select the criteria that it finds useful for the purpose of assessing a school's capacity and performance; and
- *system performance standards* that provide evidence about the success of schools, local school systems, states, and the Nation in bringing all students, leaving no one behind , to high performance standards. (p. 13)

Content standards and performance standards, operationalized through assessments, are the elements of the standards-based reforms described in the early 1990s (e.g., Smith & O'Day, 1991; Clune, 1993) that found their way into actual policies.

Content standards are broad descriptions of what students should know and be able to do in a given subject area, and are perhaps the aspect of the reforms that are best understood by the public (McLaughlin & Shepard, 1995). Following the lead of national mathematics and science education groups, states, districts, and schools began to create standards writing teams in

different subject areas to develop general statements about what their students should know and be able to do at different levels.

The origins of standards policy as a basis of instructional change were based on constructivist ideas about learning, ideas that involve more cooperative student grouping structures and more active learning in classrooms than recognized by traditional, teacher-centered, didactic models of instruction (NCTM, 1989, 1991, 2000). These standards documents (and others like them from the National Research Council, National Committee on Science Education Standards and Assessments, 1995; Rutherford & Ahlgren, 1990) called for a shift in mathematics curriculum from emphasizing computation and rote memorization of facts and procedures to emphasizing conceptual development and engaging all students in developing their mathematical reasoning power.

These standards documents incorporated a vision for instruction that

- emphasizes high expectations for all students;
- engages students in meaningful activities that enable them to construct and apply their knowledge of key concepts;
- reflects sound principles from research on how students learn, including the use of cooperative learning techniques promoting interaction and deeper understanding;
- features appropriate, ongoing use of calculators, computers, and other technologies for learning;
- is based on deep teacher understanding of subject matter; and
- makes use of ongoing support for classroom teachers, including continuing opportunities for teachers to work with one another in planning curriculum, instruction, and assessment (Weiss, 1994).

Performance standards more precisely and explicitly specified the knowledge and tasks that students must successfully complete in order to demonstrate mastery of the content standards. The latter were typically exemplified through the assessments used to measure student achievement. The breadth of change in instruction initially implied by the reforms implied equally sweeping changes in assessment practices and formidable technical obstacles.

The inclusion of performance standards within the purview of standards-based reform originated in a conception of standards as fostering new forms of instruction, and drew upon a history of assessment reform geared to higher expectations for student learning. Since content standards were to exemplify complex, higher order skills and thought processes, reformers argued that using the sorts of low-level, multiple-choice standardized assessments that had historically been used for ranking and measuring students over the years was inappropriate. Alternate assessment systems were needed, based on constructivist reform emphases on "higher order" thinking skills and a considerable body of research showing the corrupting effects that

widely used, standardized, multiple-choice assessment measures have had on such ambitious learning goals.

Resnick and Resnick (1992) influenced early conceptions of standards reform by suggesting that complex assessments should be used to drive improvements in instruction. They reviewed the historical relationship between assessment and instructional programs, and concluded that widely used multiple-choice standardized achievement tests of basic skills drive curriculum and instruction toward low-level expectations of students. As an alternative, they advocated using performance assessments to measure higher order thinking and content to drive instruction toward what they called the "thinking curriculum." In addition to measuring higher order thinking, such types of assessments (e.g., rubrics, portfolios, or student-generated projects) would be considered more "authentic" (better aligned with classroom instructional practice) than more standardized measures (Wiggins, 1989; Shepard, 1989). Second, they were theorized to be potentially more "systemically valid" (Frederiksen & Collins, 1989) and less easily corruptible than high-stakes standardized measures. This means that improved student test scores based on performance assessments are thought to reflect student learning more validly than scores based on large-scale, pencil-and-paper measures. They are thought to be less susceptible to the score pollution caused by factors such as overt test preparation that tend to be encouraged by a high-stakes testing environment (Haladyna et al., 1991).

Delivery or opportunity-to-learn (OTL) standards proved to be contentious during early conversations about standards reform. At its most general level, the concept of OTL addresses whether students have had the opportunity to study a particular topic or to learn how to solve a particular type of problem presented by an assessment (McDonnell, 1995). This concept is an important prerequisite for assessing whether comparisons of student achievement are valid, and was originated by international researchers in the early 1960s as a way to increase the validity of cross-national comparisons of student mathematics achievement. Eventually, OTL measures were refined to address classroom-specific processes, including whether teachers had taught the content needed to respond to specific items administered on the test, and whether teachers' general goals, beliefs, instructional strategies, and professional preparation provided their students with such opportunities (Schmidt & McKnight, 1995). Therefore, OTL standards and measures held promise for operationalizing and examining the nature of classroom processes.

However, the potential of OTL standards for changing how education resources were traditionally allocated made them a target in the policy development process. Advocates of standards for OTL envisioned them as a way to hold policymakers accountable for providing adequate learning opportunities to students traditionally underserved by the education system (O'Day & Smith, 1992), but others raised concerns about whether these standards were an appropriate vehicle for addressing the equity and quality problems of education (McLaughlin & Shepard, 1995; Traiman, 1993). Hot debate swirled, primarily along partisan lines, around how such standards would be defined, what their purpose and use should be, when they should be developed during the implementation process, and what the role of the federal government should be in setting them (Traiman, 1993).

Policy debates and technical problems of operationalization and measurement persisted. OTL standards, although incorporated into policy in the initial 1994 Goals 2000 legislation, were subsequently repealed in 1996. Since then, they have remained largely outside the realm of

policy, although they, along with *system performance standards*, continue to be addressed in the research community. The omission of OTL and system performance standards from the standards-based reform discussion is particularly important since current standards policies are increasingly based on high-stakes accountability assessments with consequences for schools and districts, educators, and, increasingly, students. Several lawsuits about OTL have occurred in Texas (e.g., *GI Forum*, 2000) and California (see Sandham, 2000), usually when test performance is the basis of making important decisions about students. Researchers and others are questioning the fairness of holding schools and students accountable for performance without evidence of adequate policy investment in school capacity and student learning opportunities.

The Current Context for Standards in Policy and Practice

In part as a result of the emphasis in policy on investing in content standards and relatively economical tests of achievement instead of investing in student opportunities and system performance standards, current education policy is characterized primarily by content standards and large-scale assessments serving as performance standards. Thus, standards *as the basis of accountability* is the overriding paradigm within the policy context. State standards documents tend to be lists of topics for coverage that lack explicit guidance for teachers about appropriate instructional strategies, and, indeed, do not provide enough concrete illustrations and examples to educate teachers about how to interpret those standards (Hill, 2001). Further, current assessment policies rely primarily on large-scale measures and mainly address school accountability purposes, rather than providing feedback for teachers to fine-tune their instruction. *Education Week* (2001), in a recent analysis of accountability reform in the United States, notes that 49 of the 50 states have state standards and 50 have or are in the process of developing state assessments to measure student achievement. Forty-five states generate school report cards, and 27 of those rate school performance primarily on student test scores.

Test scores do not necessarily reflect progress toward ambitious learning goals; most currently used accountability assessments are hybrids of older norm-referenced tests and newer, more expensive, criterion-referenced measures. In general, these assessments use a combination of response formats, are published by a handful of large national test publishers, and often do not align with state and local content standards (Webb, 1999; Gandal & Vranek, 2001). They represent relatively traditional approaches to measurement and tend to measure relatively lower order skills. According to Hattie, Jaeger, and Bond (1999):

> The underlying assumptions of our present major measurement models (classical and item response) seem to be grounded on Bloom's first one or two levels of knowledge and comprehension; they are more capable of modeling these two levels and less capable of modeling the higher order processing proficiencies. (p. 433)

These measures do not relate systematically to the idea of standards as the basis of instruction to support ambitious learning. According to Gandal and Vranek (2001), "Tests don't need to measure only the most rudimentary skills and knowledge, but they often do" (p. 10). Therefore, a perspective that equates standards solely with student performance on accountability measures is likely to omit certain areas of higher order thinking integral to early conceptions of

standards as part of the "thinking curriculum." Such a perspective is not likely to consider that instruction should necessarily address higher order aspects of learning like those on the higher levels of Bloom's taxonomy. Therefore, such a perspective may very well place constraints on students' opportunities to learn these higher order processes.

Looking at the ways in which local educators interpret and put into place standards policies is an important part of any examination into what happens in schools and classrooms in the name of standards. The ways in which local policymakers construct their understandings of external policies have been found to be key to the local implementation of complex reforms (Firestone, 1989; Spillane, 1998; Weatherley & Lipsky, 1977). This has been particularly well explored in the case of teachers as policymakers relative to classroom instruction (Cohen & Ball, 1990, 1999). On the other hand, school and district administrators' understandings and actions relative to policy shape the contexts in which teachers operate (Price, Ball, & Luks, 1994; Spillane, 1994). Principals' interpretations of standards policies and the ways in which they shape these interpretations into leadership practice help to define local priorities for teachers. Principals help to articulate and define those interpretations of standards that are organizationally acceptable in their schools — whether test performance is the most important thing, or whether certain instructional emphases are also important.

Other organizational factors also may play a part in the ways in which local educators construct their understandings about standards policies. For instance, there is some evidence that interpretations of standards and their implications for instruction vary depending on school performance level and capacity. Teachers have been found to interpret the implications of New York's state standards for changes in instruction differentially according to their school's conditions (McGill-Franzen & Ward, 1997). If teachers were under pressure to improve test scores (which, in this instance, were not aligned with the state standards) and they worked with little authority and internal accountability, they were not likely to consider the standards as an avenue for changing their instructional practices in significant ways.

A considerable body of research has examined the relationship between organizational factors and student achievement by studying reform processes in high-performing schools that serve very high-needs student populations. A consistent theme in this research is that the schools share an organizational focus or vision that relates directly to student learning (Lake, Hill, O'Toole, & Celio, 1999; Aldersebaes, Potter, & Hamilton, 2000; Education Trust, 1999; Haycock, 2001). Most of this research has not systematically explored the precise nature of this vision. Although there is some allusion to a vision of "ambitious teaching and learning" in the research, it is unclear whether this means a general trend of improvement on students' test scores or whether it actually addresses the teaching and learning processes related to higher order learning. In the recent research on high-performing, high-needs schools, although the visions described are clearly linked to student achievement, it is not clear whether such visions are, in fact, reductionist, that is, linked to performance on a single measure, or complex in that they involve more higher order learning processes and changes in teaching and learning.

In the research about schools that have changed instructional practice in ways that are consistent with the idea of standards as requiring high expectations for all students, organizational vision is described more clearly. In such schools, the vision is specific to teaching and learning and is expressed through a *shared technological language* indicating a common

understanding about what it means specifically for classroom practices (Jennings & Spillane, 1996; Elmore, Peterson, & McCarthey, 1996; Learning Research and Development Center [LRDC], 1998). Whether such a language develops and supports instructional change depends on the extent to which different individual understandings of policy are revealed and discussed at different levels; it also depends on the nature of the organization's discourse community (Spillane, Peterson, Prawat, Jennings, & Borman, 1996). A key aspect of clarity is knowledge about and skills related to implementing new instructional strategies and challenging curricula (Newmann & Wehlage, 1995).

The study reported here provides initial information about teachers' perspectives on elements of standards-based education in one high-performing, high-needs school district and addresses three research questions:

1. What are teachers' perspectives about key aspects of standards-based education?
2. What emphases do principals communicate in terms of the importance of student test results and the importance of instructional practices?
3. To what extent do teachers' perspectives and principals' emphases vary by school conditions (e.g., poverty and achievement level)?

METHOD

Sample, Participants, and Data Collection

This study draws on a subset of data collected in a larger study of high-needs elementary schools in one high-performing school district in a state in the Midwest. It used a two-stage stratified sampling design of schools within district. At the first stage, the state's definitions of "high-needs" and "high-performing" were enjoined to select a district that had a number of schools meeting two criteria:

1. more than 50% of the students enrolled in these schools are eligible for federal free or reduced lunch (F/RL) benefits; and
2. the proportion of students in the school who perform at the "advanced" or "proficient" level on the state accountability assessment exceeds the state average in mathematics and in reading and writing.

In developing the sample frame, data about 2000–2001 student F/RL eligibility and information about school-level performance across 1999–2000 accountability assessments in grade 3 reading, grade 4 reading and writing, and grade 5 mathematics were gathered and analyzed using data from all elementary schools in the state.

The information about school-level achievement and poverty was examined first through cross-tabulations, as illustrated in Table 1, and then in an exploration of case summaries.

Table 1. State-Level Distributions of Schools According to Poverty and Student Achievement

School Poverty	Student Achievement	
	Schools with a lower rate of proficient or advanced students than the state overall	Schools with a higher rate of proficient or advanced students than the state overall
Extreme poverty (More than 75% of students eligible for F/RL)	69	1
High poverty (Between 50.1% and 75% of students are eligible for F/RL)	118	16
Moderate poverty (Between 25.1% and 50% of students are eligible for F/RL)	123	87
Low poverty (25% or fewer of students are eligible for F/RL)	44	292

Note. A maximum of 806 different schools participated in the 1999-2000 state tests, but many schools administered those tests to small numbers of students. Those data are counted as missing cases in the state data set, and were excluded from this cross-tabulation.

The district selected for study accounted for the single high-performing, extreme poverty school in the state and six of the 16 high-performing, high poverty schools. Further, a multi-year analysis of achievement scores in this district indicated that the 1999–2000 test data continued an ongoing trend. The majority of schools in the district had demonstrated high performance over multiple years. Cross-tabulations of the district's 22 elementary schools by poverty and performance are shown in Table 2.

Table 2. Distribution of Elementary Schools in Selected District

School Poverty	Student Achievement	
	Schools with a lower rate of proficient or advanced students than the state overall	Schools with a higher rate of proficient or advanced students than the state overall
Extreme poverty (More than 75% of students eligible for F/RL)	7	1
High poverty (Between 50.1% and 75% of students are eligible for F/RL)	3	6
Moderate poverty (Between 25.1% and 50% of students are eligible for F/RL)	0	5

No elementary schools in the district enrolled fewer than 25% of students who were eligible for F/RL benefits during the 2000–2001 school year, but schools were otherwise relatively variable. The district itself is relatively small, serving approximately 17,500 students. It is located in a small industrial city, with a student population that is predominantly Hispanic (approximately 53%, many of whom are English Language Learners), and White (approximately 43%). Other ethnicities account for less than 5% of the total student population. District leadership has credited high student achievement results to increasing central oversight of curriculum, a district wide focus on literacy and student achievement, investment in professional development specific to targeting reading difficulties, and the use of frequent, ongoing student achievement data to track student progress and target assistance accordingly, in what has been documented as a coordinated district approach to the state test (e.g., Firestone & Fairman, 1998). Additionally, during the 1990s, the district had participated in several standards-oriented reform projects related to mathematics and science.

The second stage of sampling involved selecting ten elementary schools from across the range of poverty and achievement levels represented in the district. After taking into account school performance on 2000–2001 reading, writing, and mathematics tests, and calculating achievement based on three years of student language arts data and two years of mathematics data, schools were selected. These schools varied in size and poverty level, but generally had relatively experienced teachers, consistent class sizes (school averages for language arts instruction ranged between 19 and 27; for mathematics, between 20 and 28 students). Schools also spent approximately twice as much instructional time on language arts (at least two hours per day, in general) as on mathematics (approximately one hour per day), which was related to a district scheduling requirement for literacy instruction. Specific characteristics of sampled schools are provided in Appendix A.

Within each sampled school, all K–5 teachers (including regular classroom teachers, Special Education teachers, Title I teachers, and teachers of English Language Learners) were surveyed in the spring of 2001. A total of 172 teachers out of an estimated 249 teachers in selected schools completed surveys about professional development opportunities, school organization, leadership and culture, classroom practices in mathematics, reading, and writing, and teacher beliefs and attitudes about standards-related policies and practices, for a 69% response rate. Survey results were entered into an SPSS database for analyses.

Survey items used in this study included the following:

1. Six items measured teacher beliefs about key aspects of standards-based reforms in mathematics and language arts. These items included teachers' beliefs about the usefulness and quality of state and district tests of standards, about the relationship between testing and improved student learning, and about whether members of the school community hold a common opinion about what standards-based instruction looks like.

2. One item measured teacher beliefs about the role of the state test relative to school curriculum in general.

3. Six items measured teachers' perceptions of the degree to which principals emphasized test results versus different aspects of instruction in their evaluations.

4. One item measured teachers' opinions about whether administrators understand the implications of standards-based education for classrooms.

Specific items and response formats are provided in Appendix B of this report.

Analysis

Initially, descriptive statistics were used to analyze how teachers perceive different elements of standards-based reforms in these schools, and what they think their principals consider to be important. Then, several factor analyses were conducted, the results of which were used in subsequent analyses (depending on their scale reliability). Finally, Analyses of Variance were conducted to examine whether teachers' perspectives and principals' emphases varied systematically with school conditions (e.g., poverty level and achievement level).

RESULTS

Teacher Perspectives

To address the first research question, descriptive statistics were used to analyze results on seven items measuring agreement about different aspects of standards-based education. These indicated that teachers described a variety of perspectives relative to different aspects of standards-based education, and that their responses were similar for mathematics and language arts. To identify underlying constructs and to reduce the amount of data to be used in examining group differences by school conditions, two exploratory factor analyses were conducted — one for each content area. Details of the factor analyses are provided in Appendix C. The results of the factor analyses uncovered one sufficiently reliable factor for use in future analyses; the factor addressing teacher assessment of the *quality and usefulness of external tests for improving student learning to standards* (α = .7676 for language arts and α = .7568 for math). These results provide empirical validation about an underlying construct addressing teachers' perspectives about the effectiveness and quality of external tests.

Perspectives about Testing, Test Use, and Learning Opportunities

Chart 1 displays teachers' perspectives about testing, test use, and learning opportunities relative to mathematics and language arts standards.

As shown, these teachers generally agreed that the use of test results has improved student learning; more than three-quarters of respondents indicated that using test results to guide instruction had improved student learning across both content areas. Similarly, the group generally agreed that additional learning time is provided for those students that are not proficient on the state test in mathematics and in language arts. Both findings are consistent with descriptions of how high-performing, high-needs schools and districts have used achievement

data to strengthen the performance of students who are most at risk for failure (e.g., Education Trust, 1999; Ragland, Asera, & Johnson, 1999).

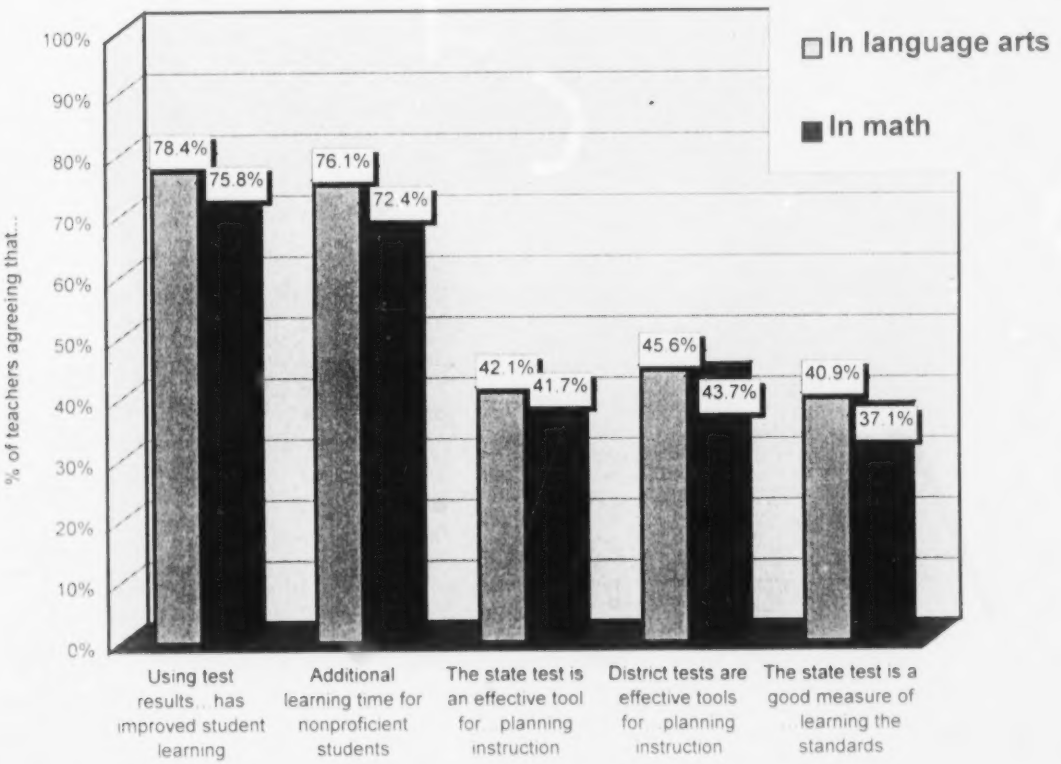

Chart 1. Teacher perspectives on testing and test use

Although teachers describe the use of test results to guide teaching as being, in general, a force for improvement of student learning, they are much less positive about the usefulness of district and state tests for identifying students' needs and thus helping them plan instruction. This seems somewhat contradictory, but patterns in the qualitative data gathered indicate that the district has adopted a practice of using test results (described generally in the first item) to guide teaching so that classroom content coverage is aligned with standards/or the test (Snow-Renner & Reichardt, 2001). In other words, test results are used to identify *content* that is tested but not taught in particular programs. The items about the effectiveness of district and state tests in this study were specific to a different use, namely, as *diagnostic* tools for identifying *students' needs* and helping teachers to plan instruction accordingly.

As might be expected, slightly more teachers viewed the district tests as being more effective diagnostic and planning tools than the state test. Approximately 45% of teachers agreed

that district tests are effective for identifying students' needs and planning instruction, compared to about 42% who judged the state test as being good for the same purpose.

In general, a majority of teachers did not agree that the state test is a good measure of student learning of the standards; across both content areas, approximately 60% of teachers said that it is not. This indicates that the teachers in these schools perceive a disjunction between the test and the standards and that they do not necessarily equate student attainment of the standards solely with how well they do on the state accountability measure. They seem to have a conception of something more. This may entail a broader conception of content than is covered by mathematics and reading, the two areas in which the state has conducted the most extensive testing in elementary grades.

Perspectives about the Curriculum and the State Test

To examine the relationship between the test and curriculum, teachers were again asked to rate their agreement with the following statement on a four-point Likert scale, with response options ranging from "strongly disagree" to "strongly agree."

Emphasis on the state test has narrowed the curriculum and reduced time spent on content areas that are not tested.

Chart 2 displays an illustration of response frequencies.

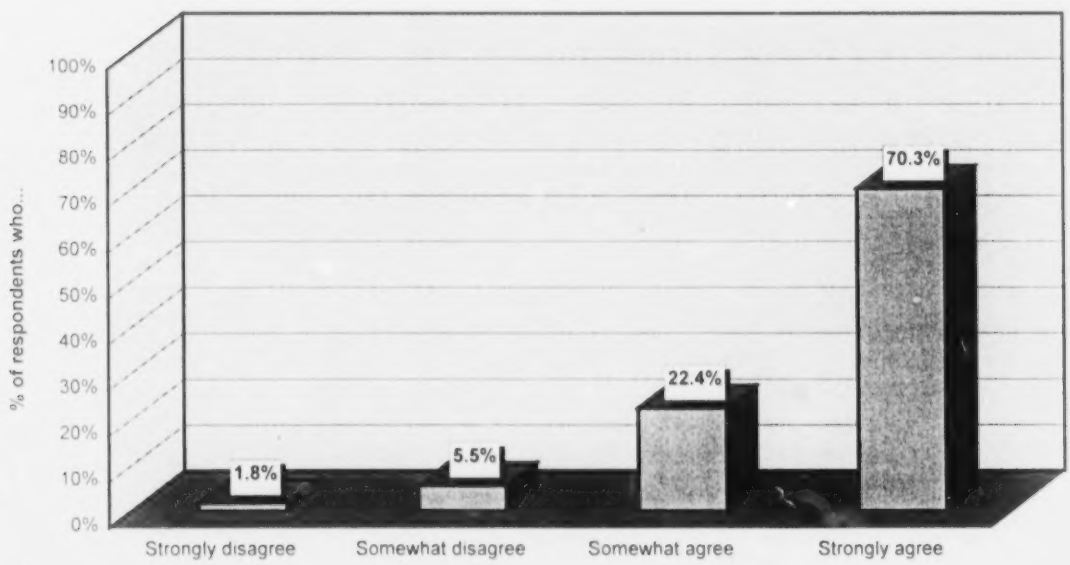

Chart 2. Teacher perspectives on the relationship between the state test and curriculum

Overwhelmingly, respondents (93%) agree (and more than 70% strongly agree) that emphasis on the state test has narrowed the curriculum and that it has reduced the time they spend on content areas that aren't tested. It is apparent that some things are being crowded out of

the curriculum because of the emphasis on the state test, although these data do not shed light on what those things may be.

Perspectives about Instruction Based on Standards

The research indicates that such a language exists in schools that have changed instruction in ways consistent with standards as a basis of instruction. There are clearly defined standards for what constitutes good teaching practice that are explicit, widely discussed, and modeled wherever possible (Learning Research and Development Center, 1998). Further patterns in the data raise questions about whether these teachers share a common technological language about what standards-based education looks like in mathematics and science classrooms. Chart 3 provides an illustration of teachers' responses to an item measuring the common understanding of what standards-based teaching means within these schools. More than 80% of the elementary teachers surveyed agreed that in their schools, people held different opinions about how to practice standards-based teaching in language arts. The percentage was slightly lower for mathematics; approximately 67% of respondents agreed that this was the case. Although approximately 60% of respondents voiced moderate agreement that people hold different opinions in their schools, almost 20% of respondents voiced strong agreement about the disagreement in their sites — and this held true across content areas.

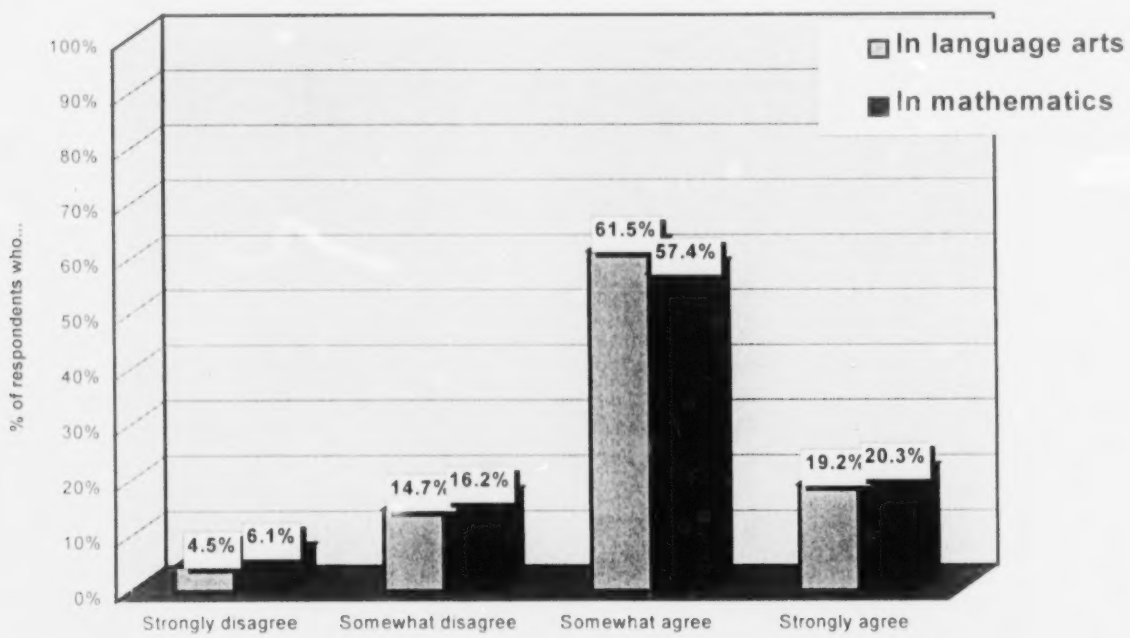

Chart 3. Teacher perspectives about common understandings of standards-based teaching

Principals' Emphases

District and school leaders send important messages to teachers about appropriate actions and what it takes to implement a reform (Price et al., 1994). In particular, the principal's supervisory role serves as an important guide to teachers about what is expected of their classroom practice in relationship to overall school improvement efforts. Classroom observations related to teacher evaluation are one way that principals who are effective change agents communicate and promote a shared vision within the school. These vision-building and educational leadership tasks have been found to have the most positive effects on instruction (Liberman, Falk, & Alexander, 1994; Rosenholtz, 1989; Haynes, 1998; Puma et al., 1997; Shields, Knapp, & Wechsler, 1995; Stringfield, Datnow, & Ross, 2000).

In this study, teachers were asked to rate a number of specific elements that they thought their principals considered important in evaluating teacher performance. These elements were linked to either instructional practices or student test performance, and included the following:

- Teachers' use of specific instructional strategies, demonstration of content knowledge, individualization of instruction for different students, and student engagement in learning activities,
- How well students do on the state test and other tests and how much teacher instruction focuses on raising student test scores.

A confirmatory factor analysis of these 6 elements verified that they represent two independent factors (refer to Appendix C).

Charts 4 and 5 provide illustrations of responses to items on the instruction and testing scales. These charts indicate that, overall, teachers perceive their principals to have a relatively balanced emphasis on instruction and on student achievement. Although it is apparent that almost everything is considered important, it is also apparent that principals consider instructional practices to be slightly more important than test scores. More than 97% of teachers agree that principals consider their use of specific instructional strategies, their content knowledge, their strategies for individualizing instruction for different students in their classes and how engaged students are in learning activities as very important or moderately important. Very few teachers report that they don't know if these things are important to principals. More than 81% of teachers say that test results are very important or moderately important to their principals, and 88% say their principals consider their instruction focused on raising test scores as very important or moderately important, while approximately 8–10% of teachers report that they don't know how important these things are to their principals.

These response patterns indicate while that these elementary school principals communicate that a primary emphasis for teachers is on the *quality of instructional practice* in their classrooms, they also pay a good deal of attention to their students' test scores.

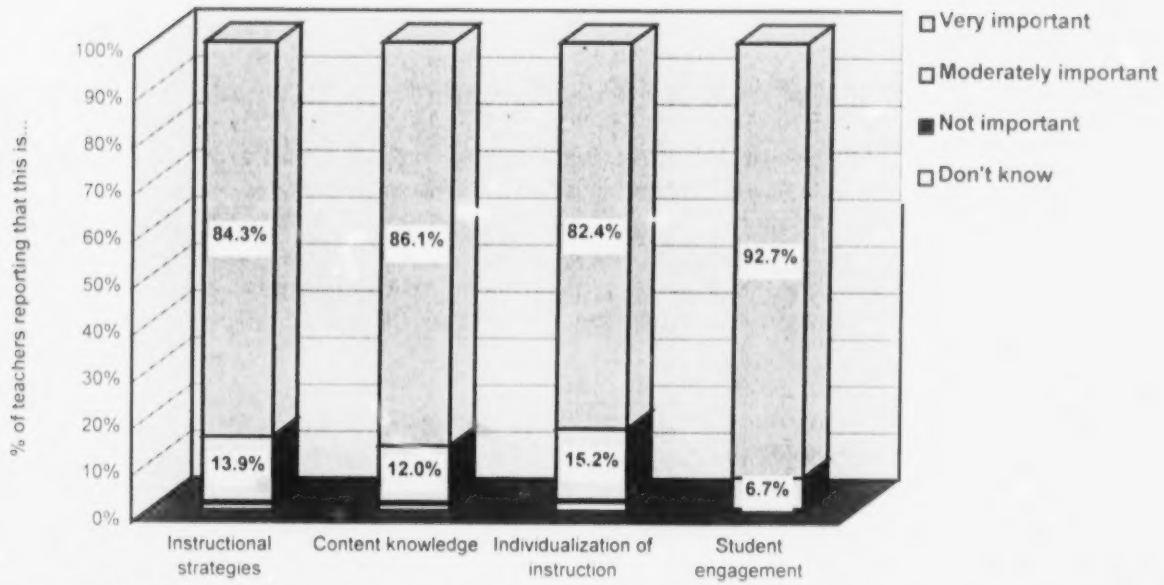

Chart 4. Teacher perceptions of the importance of instructional practice to principals

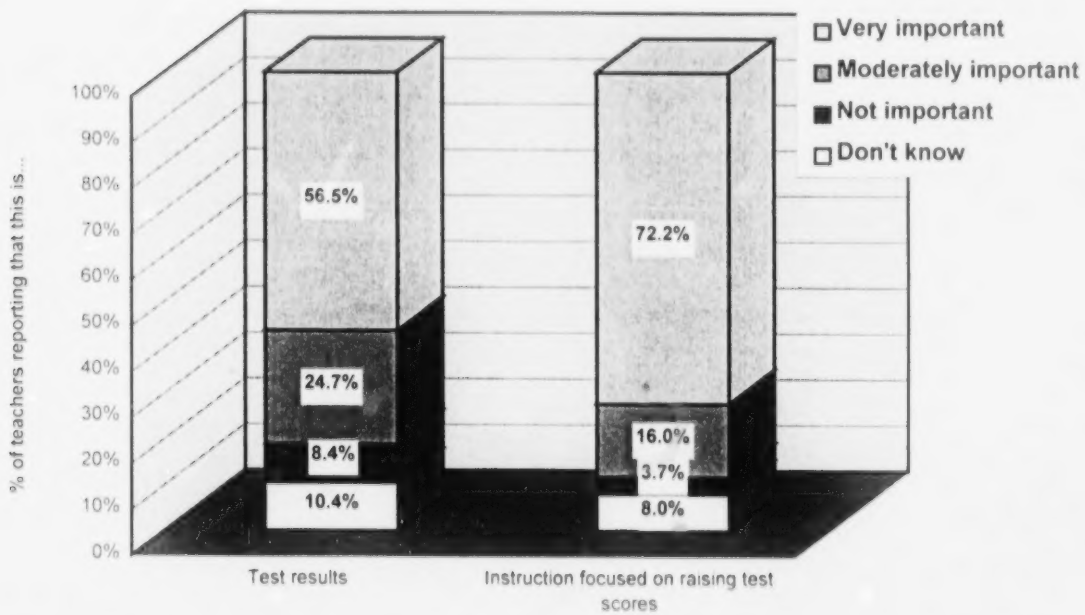

Chart 5. Teacher perceptions of the importance of their students' test results to principals

Further, teachers express considerable confidence in the knowledge of their administrators about what standards should look like in the classroom, indicated by their responses to the survey item

Administrators in my school don't understand what standards mean for classroom instruction.

Responses indicate that more than 86% of teachers disagree with the item. Chart 6 displays response frequencies on this item.

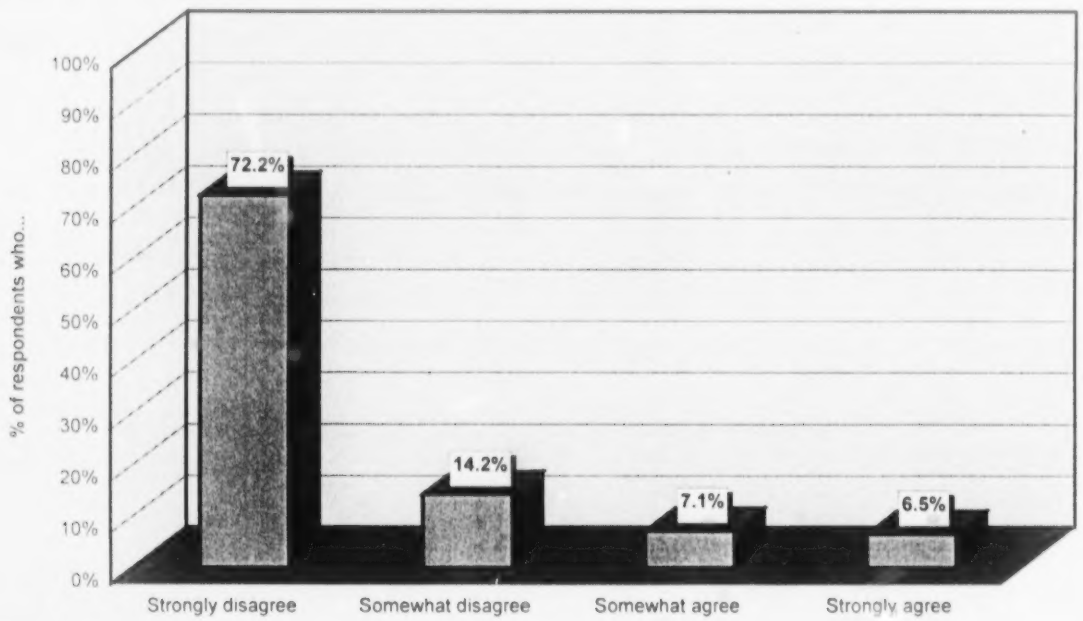

Chart 6. Teacher perspectives about administrators' understanding of standards

The high confidence that teachers express in administrators' understanding of what standards imply for instruction, and the considerable emphasis on instructional practices that principals communicate to teachers, are evident in these data. However, the survey does not provide substantive information about the extent to which quality instruction, as it is understood in these schools, is congruent with the idea of standards-based instruction as defined in early standards documents. Such an exploration is beyond the scope of this particular study and entails a detailed examination of the ways in which standards are interpreted relative to classroom practice, as well as a more comprehensive analysis of classroom practices.

Differences by School Conditions

Analyses of variance were used to examine whether teachers' perspectives on standards and principals' emphases on test results and instruction varied across school conditions (the

convention of 5% significance level was observed throughout). School conditions had four levels:

1. Extremely-high-poverty schools with high achievement (teacher n = 13)
2. Extremely-high-poverty schools with low achievement (teacher n = 32)
3. High-poverty schools with high achievement (teacher n = 47)
4. Moderate-poverty schools with high achievement (teacher n = 56)

The dependent variables in the analyses included the following from the factor analysis:

- Factor scores addressing teachers' assessment of the quality and usefulness of tests for improving student learning to standards.
- Scale scores measuring principals' emphases — on instruction and on test results.

Also used as dependent variables were teacher ratings of agreement with statements about

- narrowing of the curriculum in response to emphases on the state test,
- common opinions in the school about how to practice standards-based teaching,
- use of test results relative to improving student learning, and
- provision of additional learning time for students who are nonproficient on the state test.

Differences in Teacher Perspectives

No significant differences were found across school conditions in how teachers assessed the quality and usefulness of external tests for improving students' learning to standards in mathematics or in language arts. Teachers exhibited relatively low agreement with the idea that the state test is a good measure of students learning the standards, as well as the idea that state and district tests helped them to identify students' needs and to plan instruction accordingly.

In addition, there were no significant differences in the extent to which teachers in different types of schools reported that emphasis on the state test had narrowed the curriculum. Similarly, teachers in different types of schools did not exhibit significant differences in the extent to which a shared vision exists about what standards mean for classroom instruction. It appears that regardless of the poverty and achievement level of the schools they teach in, teachers do not generally share a common opinion about what standards mean for instruction.

Significant differences emerged, however, in other areas. First, the extent to which teachers felt that the use of test results to guide teaching had improved student learning in the

school varied significantly by school conditions. This was the case both in language arts (F = 7.032, df = 3, 158, p <.001) and in mathematics (F = 9.011, df = 3, 149, p <.001). Post-hoc tests for differences indicated that teachers in moderate poverty schools are significantly less likely than teachers in all other schools to feel that the use of test results to guide teaching has improved student learning in either language arts or mathematics. Chart D-1 in Appendix D provides a report of mean responses by school condition.

Second, results indicate that the provision of additional learning time in mathematics and in language arts for students who are not proficient on the state test varied significantly by school conditions as well (in mathematics, F = 3.746, df = 3, 147, p = .012; in language arts, F = 6.677, df = 3, 144, p <.001).

In language arts, post-hoc tests for differences indicate that teachers in schools with moderate poverty levels agree to a significantly lesser extent that additional learning time is provided for students who are nonproficient than teachers in high-poverty, high-performing schools (p = .028) and teachers in extreme poverty, high-performing sites (p = .091). These findings may reflect the staging of the district's rollout of its intervention program in literacy, which has targeted higher poverty schools prior to lower poverty schools. The district's five lowest poverty elementary schools are scheduled to receive this intervention in the 2001–2002 school year.

In the area of mathematics, post-hoc tests revealed two significant differences. First, in extremely-high-poverty schools, there is a significant difference between high-performing and low-performing schools. Teachers in high-performing schools are significantly more likely to agree that nonproficient students receive additional learning time in mathematics than teachers in low-performing schools (p = .002). Further, among high-performing schools of moderate, high, and extreme poverty, significant differences exist. Teachers in extreme-poverty, high-performing schools and teachers in high-poverty, high-performing schools are significantly more likely to report additional learning time for nonproficient students in mathematics than are teachers in schools of only moderate poverty (p <.001 and p = .005, respectively). Chart D-2 in Appendix D provides illustrations of mean responses about the provision of extra learning time in language arts and in mathematics for nonproficient students, by school condition.

Differences in Principals' Emphases

Analyses of variances of differences in teachers' perceptions of principals' emphases on instruction or on test results by school conditions provided mixed results. Principals' emphasis on test results was not significantly different by school type (F = 1.76, df = 3,149, p = .157). However, principals' emphasis on instruction was significantly different (F = 4.299, df = 3, 160, p = .006) by school condition. Post-hoc tests of difference indicated that the only significant difference is between principals in high-performing, extreme-poverty sites who communicate a significantly greater emphasis on instruction than their counterparts in low-performing, extreme-poverty sites (p = .022). Chart D-3 in Appendix D provides an illustration of mean responses about principals' emphases on test results and on instruction by school conditions. Table 3 provides a summary of significant differences in teachers' perspectives and principals' emphases across different school conditions.

Table 3. Summary of Significant Differences by School Conditions ($p<.05$)

Differences in ...	Specific variables or scales	Group differences
Teachers' perspectives	Test results have improved student learning in language arts and in mathematics	Teachers in moderate-poverty schools are significantly less likely to agree than teachers in any other schools.
	Additional learning time is provided for students who are not proficient on the state test in language arts	Teachers in moderate-poverty schools are significantly less likely to agree than teachers in high-poverty, high-performing schools
	Additional learning time is provided for students who are not proficient on the state test in mathematics	Teachers in extreme-poverty, high-performing schools are significantly more likely to agree than teachers in extreme-poverty, low-performing schools Teachers in extreme-poverty, high-performing schools and in high-poverty, high-achieving schools are significantly more likely to agree than teachers in moderate-poverty schools.
Principals' communicated emphases	Emphasis on instruction (e.g., teachers' content knowledge, use of specific instructional strategies, individualization for different students, and student engagement.	Teachers in extreme-poverty, high-performing schools rate their principals significantly higher on their instructional emphasis than teachers in extreme-poverty, low-performing schools

DISCUSSION

This study provides initial information about elementary teachers' perspectives on key aspects of standards-based education. Results indicate that teachers respond to particular aspects of standards reforms in similar ways across language arts and mathematics content. Most teachers say that using test results to inform teaching has improved student learning in their schools and that additional learning time is provided for students who are not proficient on the state test, although no data were collected to determine how this time was used. Teachers also raise questions about the usefulness of externally mandated tests for helping them plan instruction for individual students. Further, in terms of instruction related to standards, these teachers do not describe their schools as organizations in which broad agreement exists about how to practice standards-based instruction in the classroom. However, they do say that their principals consider their instructional practices to be very important for judging their teaching performance — and that instructional practice is slightly more important to principals than their students' test scores. Teachers also express confidence in their principals' knowledge about standards, particularly in terms of the implications of standards for instructional practice.

The data explored in the study illuminate some key differences in teacher perspectives by the school variables of achievement and poverty level. First, teachers in the most affluent schools in the sample are significantly less positive than teachers in other schools, regardless of their achievement levels, about the use of test results for improving student learning in mathematics and in language arts. Second, teachers in these schools report providing additional learning time for nonproficient students in mathematics and in language arts significantly less than do teachers in other high-performing sites that serve poorer populations. This may indicate that higher achievement in the poorer schools is partially due to targeting additional time to students who are

most at-risk. Finally, teachers in high-performing schools that serve extremely poor student populations report providing additional time in mathematics to nonproficient students significantly more than teachers in low-performing schools that serve extremely poor student populations.

The data from this study focused primarily on teachers' perspectives about different aspects of standards-based reform. Classroom practices were not observed. Nevertheless, the data suggest that teachers are cognizant of (a) the differences between standards *per se* and the state test; (b) different uses for test results, such as targeting remediation opportunities for low-achieving students and examining the mismatches between the enacted curriculum and what is tested; and (c) the dependence of utility of external tests on the uses to which results are put.

Thompson (2001) has described the current standards policy context in terms similar to the distinction described at the beginning of this paper between *standards as the basis of instruction* and *standards as the basis of accountability*. He has contrasted one perspective as "authentic, standards-based reform," which is characterized by its emphasis on improving the quality of instructional content and practice and its concern for equity, with its "evil twin" of "high-stakes, standardized, test-based reform" (p. 358), in which equity issues tend to be subverted, and instruction may very well be shaped in ways that are antithetical to the early intent of standards reformers. Thompson states that "what gets lost when teachers and students are pressured to make students better test-takers is precisely the rich, high-level teaching and learning that authentic, standards-based reform aims to promote in all classrooms and for all students" (p. 358).

The results of this study indicate that the extreme contrasts in standards perspectives that have been painted by Thompson (2001) are overly simplistic. The actual picture is much more complex and is likely to be influenced by a variety of things, including the quality and number of tests given, how results are used, and how teachers are supported in learning how to use different diagnostic and instructional tools to improve student learning processes. In the district studied, leaders have tried to make the results of assessments useful not only for accountability purposes and comparisons of school performance, but also for teachers to plan their instruction, both in terms of its alignment with standards and to diagnose individual students' learning needs. Whether this is actually possible — whether assessments can actually serve these multiple purposes equally well — has been a matter of disagreement among policy makers and assessment experts for some time (e.g., Stiggins, 1992; McDonnell, 1994). It remains particularly problematic in a high-stakes policy environment that provides teachers with incentives to maximize scores independent of their validity for instructional uses.

The data suggest that teachers do not, in general, share common understandings about what standards mean for language arts or mathematics instruction in these schools. The data also highlight the fact that teachers and principals consider high quality instruction to be very important, if only slightly more so than test scores. However, it is not possible to determine from these data if "high quality instruction" is interpreted in these schools as something that is congruent with the early ideas about instruction reflected in standards documents (NCTM, 1989; Weiss, 1994), or as high proportions of students being classified as proficient or advanced on the state test.

A related question is the principal's role in instructional change. A number of research studies document the importance of the principal's role in changing instruction and raising student achievement (Davidson & Taylor, 1998). In the schools studied, teachers describe principals as being very knowledgeable about what the implications of standards are for classroom instruction and they also describe principals as attempting to strike a balance between emphasizing accountability and instructional quality.

REFERENCES

Aldersebaes, I., Potter, J., & Hamilton, N. (Dec 1, 2000). *Programs don't, people do: Insights into schoolwide change*. Portland, OR: Northwest Regional Educational Laboratory.

Clune, W. H. (1993). The best path to systemic educational policy: Standard/centralized or differentiated/decentralized? *Educational Evaluation and Policy Analysis, 3*(15), 233–254.

Cohen, D. K., & Ball, D. (1990). Relations between policy and practice: A commentary. *Educational Evaluation and Policy Analysis, 12*(3), 331–338.

Cohen, D. K., & Ball, D. L. (1999, June). *Instruction, capacity, and improvement* (CPRE Research Report Series RR-43). Philadelphia: Consortium for Policy Research in Education.

Conference Report on H.R. 1804, Goals 2000: Educate America Act. (1994). *Congressional Record, 140*(32), H1625.

Davidson, B. M., & Taylor, D. L. (1998, April). *The effects of principal succession in an accelerated school*. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.

Education Trust. (1999). Dispelling the myth: High poverty schools exceeding expectations. Washington, DC: Author.

Education Week. (2001). A better balance: Standards, tests, and the tools to success. Quality Counts 2001 [Special Issue]. *Education Week, 1*(11).

Elmore, R., Peterson, P., & McCarthey, S. (1996). *Restructuring in the classroom: Teaching, learning, and school reform*. San Francisco, CA: Jossey-Bass.

Firestone, W. A. (1989). Educational policy as an ecology of games. *Educational Researcher, 18*(7), 18–24.

Firestone, W. A., & Fairman, J. C. (1998, April). *The district role in state assessment policy: An exploratory study*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Firestone, W. A., Mayrowetz, D., & Fairman, J. (1998). Performance-based assessment and instructional change: The effects of testing in Maine and Maryland. *Educational Evaluation and Policy Analysis, 20*(3), 95–113.

Frederiksen, J. R., & Collins, A. (1989, December). A systems approach to educational testing. *Educational Researcher, 18*(9), 27–32.

Gandal, M., & Vranek, J. (2001, September). Standards: Here today, here tomorrow. *Educational Leadership, 59*(1), 6–13.

GI Forum v. Texas Education Agency, No. SA-97-CA-1278-EP (GI Forum filed Jan. 7, 2000)

Haladyna, T. M., Nolen, S. B., & Haas, N.S. (1991). Raising standardized achievement test scores and the origin of test score pollution. *Educational Researcher 5*(20), 2–7.

Hattie, J., Jaeger, R. M., & Bond, L. (1999). Persistent methodological questions in educational testing. In A. Iran-Hejad & P. D. Pearson (Eds.), *Review of research in education, volume 24* (pp. 393–446). Washington, DC: American Educational Research Association.

Haycock, K. (2001). Closing the achievement gap. *Educational Leadership*, 6–11.

Haynes, N. M. (1998). Lessons learned. *Journal of Education for Students Placed at Risk, 3*(1), 87–99.

Hill, H. C. (2001). Policy is not enough: Language and the interpretation of state standards. *American Educational Research Journal* (38), 289–318.

Jennings, N. E., & Spillane, J. P. (1996). State reform and local capacity: Encouraging ambitious instruction for all and local decision making. *Journal of Education Policy, 11*(4), 465–482.

Lake, R. J., Hill, P. T., O'Toole, L., & Celio, M. B. (1999). *Making standards work: Active voices, focused learning*. Seattle, WA: University of Washington, Center on Reinventing Public Education.

Learning Research and Development Center. (1998, March). *High performance learning communities project final report*. Pittsburgh, PA: University of Pittsburgh.

Liberman, A., Falk, B., & Alexander, L. (1994). *A culture in the making: Leadership in learner-centered schools*. New York: National Center for Restructuring Education, Schools and Teaching, Teachers College.

McDonnell, L. M. (1994). *Policymakers' views of student assessment*. Santa Monica, CA: RAND.

McDonnell, L. M. (1995). OTL as a research concept. *Educational Evaluation and Policy Analysis, 17*(3), 305–322.

McGill-Franzen, A., & Ward, N. (1997, April). *Teachers' use of new standards, frameworks, and assessments for English language arts and social studies: Local cases of New York State primary grade teachers*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

McLaughlin, M. W., & Shepard, L. A. (1995). *Improving education through standards-based reform*. Stanford, CA: The National Academy of Education.

McNeil, L. M. (2000). Creating new inequalities: Contradictions of reform. *Phi Delta Kappa, 81*(10), 728–734.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

National Council on Education Standards and Testing. (1992). *Raising standards for American education*. Washington, DC: U.S. Government Printing Office.

National Research Council, National Committee on Science Education Standards and Assessment. (1995). *National science education standards*. Washington, DC: National Academy Press.

Newmann, F. M., & Wehlage, G. G. (1995). *Successful school restructuring*. Washington, DC: American Federation of Teachers.

O'Day, J. A., & Smith, M. S. (1992). Systemic reform and educational opportunity. In S. H. Fuhrman (Ed.), *Designing coherent education policy* (pp. 250–312). San Francisco, CA: Jossey-Bass.

Price, J., Ball, D., & Luks, S. (1994, April). *Marshalling resources for reform: The role of district administrators*. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.

Puma, M. J, Karweit, N., Price, C., Ricciuti, A., Thompson, W., & Vaden-Kiernan, M. (1997). *Prospects: Student outcomes. Final report*. Washington, DC: Department of Education, Planning and Evaluation Service.

Ragland, M. A., Asera, R., & Johnson, J. F. (1999). *Urgency, responsibility, efficacy: Preliminary findings of a study of high-performing Texas school districts*. Austin, TX: The University of Texas, the Charles A. Dana Center. Retrieved August 2, 2000, from http://www.starcenter.org

Resnick, L. B., & Resnick, D. P. (1992). Assessing the thinking curriculum: New tools for educational reform. In B. R. Gifford & M. C. O'Connor (Eds.), *Changing assessments: Alternative views of aptitude, achievement, and instruction* (pp. 37–75). Boston: Kluwer.

Rosenholtz, S. (1989). *Teacher's workplace: The social organization of schools*. New York: Longman.

Rutherford, F., J., & Ahlgren, A. (1990). *Science for all Americans*. Arlington, VA: National Science Teachers' Association.

Sandham, J. L. (2000, May 24). Calif. schools lack basics, suit alleges. *Education Week 19*(37), 1, 29. Retrieved January 3, 2001, from http://www.edweek.org/ew/ew_printstory.cfm?slug=37aclu.h19.

Schmidt, W. H., & McKnight, C. C. (1995). Surveying educational opportunity in mathematics and science: An international perspective. *Educational Evaluation and Policy Analysis, 17*(3), 337–354.

Shepard, L.A. (1989). Why we need better assessments. *Educational Leadership, 46*(7), 4–9.

Shields, P. M., Knapp, M. S., & Wechsler, M. E. (1995). *Improving schools from the bottom up: From effective schools to restructuring. Summary volume*. Washington, DC: U.S. Department of Education.

Simon, E., Passantino, C., & Foley, E. (1998). *Making sense of standards: Children achieving and changing instructional practice*. Rutgers University: Consortium for Policy Research in Education.

Smith, M., & O'Day, J. (1991). Systemic school reform. In S. Fuhrman & B. Malen (Eds.), *The politics of curriculum and testing* (pp. 233–267). New York: The Falmer Press.

Spillane, J. P. (1994). How districts mediate between state policy and teachers' practice. In R. F. Elmore & S. H. Fuhrman (Eds.), *The governance of curriculum* (pp. 167–185). Alexandria, VA: Association for Supervision and Curriculum Development.

Spillane, J. P. (1998). State policy and the non-monolithic nature of the local school district: Organizational and professional considerations. *American Educational Research Journal, 35*(1), 33–63.

Spillane, J. P., Peterson, P.L., Prawat, R. S., Jennings, N. E., & Borman, J. (1996). Exploring policy and practice relations: A teaching and learning perspective. *Journal of Education Policy, 11*(4), 431–440.

Stiggins, R. J. (1992, June). *Two disciplines of educational assessment*. Paper presented at the Education Commission of the States Assessment Conference, Boulder, CO.

Stringfield, S., Datnow, A., & Ross, S. M. (2000). *Scaling up school restructuring in multicultural, multilingual contexts: Early observations from Sunland County*. Retrieved September 20, 2000, from http://crede.ucsu.edu/Programs/Program5/Project

Thompson, S. (2001, January). The authentic standards movement and its evil twin. *Phi Delta Kappan*, 358–362.

Traiman, S. L. (1993). *The debate on opportunity-to-learn standards*. Washington, DC: National Governors' Association.

Weatherley, R., & Lipsky, M. (1977). Street-level bureaucrats and institutional innovation: Implementing special-education reform. *Harvard Educational Review 47*(2), 171–197.

Webb, N. L. (1999). *Alignment of science and mathematics standards and assessments in four states*. [Research monograph #18]. Madison, WI: University of Wisconsin, National Institute for Science Education.

Weiss, I. R. (1994). *A profile of mathematics and science education in the United States*. Chapel Hill, NC: Horizon Research.

Wiggins, G. (1989, April). Teaching to the (authentic) test. *Educational Leadership, 46*(7), 41–47.

APPENDIX A. DESCRIPTIVE INFORMATION ABOUT SAMPLED ELEMENTARY SCHOOLS

	Extreme Poverty (More than 75% F/RL)			High Poverty (51-75% F/RL)			Moderate Poverty (25-50% F/RL)			
School	A	B	C	D	E	F	G	H	I	J
2000-01 Enrollment	458	377	259	264	284	449	454	489	522	486
2000-01 F/RL	80%	87%	86%	70%	67%	58%	47%	43%	36%	26%
Teacher N	23	15	16	16	20	13	15	19	26	9
School performed above state in language arts		x			x	x	x	x	x	x
School performed above state in math		x		x	x	x	x	x	x	x
Average years of teacher experience	14.9	14.1	9.1	15.1	14.2	14.1	20.2	15.1	12.8	14.8
Time spent per day in language arts instruction	2 hours, 26 minutes	2 hours, 45 minutes	2 hours, 23 minutes	1 hour, 50 minutes	2 hours, 24 minutes	2 hours, 13 minutes	2 hours, 37 minutes	2 hours, 13 minutes	2 hours, 2 minutes	2 hours, 5 minutes
Time spent per day in mathematics instruction	1 hour, 10 minutes	1 hour, 3 minutes	1 hour, 18 minutes	52 minutes	1 hour, 4 minutes	1 hour, 4 minutes	1 hour, 13 minutes	56 minutes	59 minutes	1 hour, 5 minutes
Average class size in language arts	22	21	20	19	19	27	24	25	24	22
Average class size in mathematics	22	21	20	21	20	25	21	28	24	28

APPENDIX B. SURVEY ITEMS

1. *Teacher agreement with key aspects of standards-based reforms in mathematics and in language arts.*
Response options were on a 4-point Likert scale ranging from "strongly disagree" to "strongly agree." Separate responses were required for language arts and for mathematics.

To what extent do you agree or disagree with each of the following statements in language arts and mathematics?

- Using test results to guide teaching has improved student learning in my school.
- The state test is a good measure of student learning of the standards.
- In my school, there are different opinions about how to practice standards-based teaching.
- Additional learning time is provided for students who are not proficient on the state test.
- The state test is an effective tool for identifying student needs and planning instruction.
- District tests are effective tools for identifying student needs and planning instruction.

2. *Teacher agreement with the role of the state test relative to curriculum in general.*
Response options were on a 4-point Likert scale ranging from "strongly disagree" to "strongly agree."

Indicate the extent to which you agree or disagree with each of the following statements about your school.

- Emphasis on the state test has narrowed the curriculum and reduced time spent on content areas that are not tested.

3. *Teacher perceptions of the importance with which principals emphasized test results or different aspects of instruction in their evaluations.*
Response options were on a four point scale rating each of the following as "not important", "moderately important", "highly important", or "don't know."

How important are the following things to your principal in evaluating your teaching?

- How well your students do on the state test and other tests
- Your use of specific instructional strategies
- Your demonstration of content knowledge
- How you individualize instruction for different students in your class
- How engaged your students are in learning activities
- How much your instruction focuses on raising student test scores

4. *Teachers' opinions about whether administrators understand the implications of standards-based education for instruction*

Response options were on a 4-point Likert scale ranging from "strongly disagree" to "strongly agree."

Indicate the extent to which you agree or disagree with each of the following statements about your school.

- Administrators in my school don't understand what standards mean for classroom instruction.

APPENDIX C. DESCRIPTION OF FACTOR ANALYSES EXAMINING TEACHERS' PERSPECTIVES AND PRINCIPALS' EMPHASES

TEACHERS' PERSPECTIVES

Two exploratory factor analyses were conducted — each using the following seven items:

5. Using test results to guide teaching has improved student learning in my school.
6. The state test is a good measure of student learning of the standards.
7. Some students can simply never reach high standards in this content area.
8. In my school, there are different opinions about how to practice standards-based teaching.
9. Additional learning time is provided for students who are not proficient on the state test.
10. The state test is an effective tool for identifying student needs and planning instruction.
11. District tests are effective tools for identifying student needs and planning instruction.

All seven items about standards specific to the content area that they examined were used in each factor analysis. Both factor analyses were conducted using the Principal Component Analysis extraction method (with the criterion that Eigenvalues , 1) and an orthogonal rotation. Results of the analyses were roughly parallel. Three overall factors emerged for teachers' perspectives about both language arts and mathematics, accounting for approximately 70.5% of total scale variance in each content area. General descriptions of these overall factors are provided below.

- *Factor I – Teacher assessment of the quality/usefulness of tests for improving student learning to standards* — Addresses the usefulness of state and district tests for identifying students' needs and planning instruction, as well as the quality of the state test as a measure of learning the standards. Reliability analyses of these three-item scales, using the coefficient Alpha method, indicated acceptable reliability for use of these factors in subsequent ANOVAs (, = .7676 for language arts and , = .7568 for math).

- *Factor II – Teachers' beliefs about students' capability and targeting additional learning opportunities for low-performing students*— Addresses beliefs about students' capabilities for learning to high standards and the relationship between use of test results and student learning, including the

provision of additional learning time for students who are not proficient on the state test. Reliability analyses of these three-item scales indicated that the scales were not sufficiently reliable for use in subsequent ANOVAs; therefore differences were assessed on individual items (α = .4361 for language arts and α = .4741 for math).

- *Factor III — Common understandings of standards in the classroom* — within the school, an agreement about how to practice standards-based teaching. Only the single item addressing the level of agreement loaded on this factor; therefore that item was used for subsequent ANOVAs by school conditions.

Specific items and factor loadings are summarized in Table C-1.

Table C-1. Factors and Variable Loadings related to Standards Perspectives in Language Arts and Mathematics

Items loading on factor	*Factor I — Teacher assessment of the quality/usefulness of tests for improving student learning to standards*		*Factor II — Teacher beliefs about student capability and targeting additional learning opportunities for low-performing students*		*Factor III — Common understanding of standards in the classroom*	
	factor loading in LA	factor loading in math	factor loading in LA	factor loading in math	factor loading in LA	factor loading in math
The state test is effective for identifying student needs...	.847	.831				
The state test is a good measure of learning the standards...	.793	.803				
District tests are effective for identifying student needs...	.755	.748				
All students can reach high standards*...			.766	.767		
Using test results to guide teaching has improved student learning...			.637	.648		
Additional learning time is provided...			.603	.623		
In my school, we agree about how to practice standards-based teaching*					.937	.918

* Polarity on items was reversed before factor analysis and calculation of regression factor scores.

PRINCIPALS' EMPHASES

In order to verify empirically whether underlying factors related to testing and instruction, all "don't know" responses were treated as missing cases and a confirmatory factor analysis was conducted using the following six items, the Principal Component Analysis extraction method (with the criterion that Eigenvalues , 1), and an orthogonal rotation.

How important are the following things to your principal in evaluating your teaching? (Response options: 1 = not important, 2 = moderately important, and 3 = very important)

12. How well your students do on the state test and other tests

13. Your use of specific instructional strategies

14. Your demonstration of content knowledge

15. How you individualize instruction for different students in your class

16. How engaged your students are in learning activities

17. How much your instruction focuses on raising student test scores

Two factors were derived, accounting for 68.9% of the total variance in the scale variables:

- *Factor I — Emphasis on instruction*, characterized by teachers' demonstration of content knowledge, individualization of instruction for different students in the class, teachers' use of various instructional strategies, and levels of student engagement in the class.

- *Factor II — Emphasis on test results*, characterized by an emphasis on student test scores and teacher instruction focused on raising student test scores.

These factors were used to organize and develop scale scores for use in ANOVAs of differences in principals' emphasis by school condition (Reliability analyses using the coefficient Alpha method indicated acceptable reliability for these scales, with , = .8006 for the four-item *instruction* scale and , = .7362 for the two-item *testing* scale). They were also used to organize the exploration of frequency data about how teachers perceive principals' perceptions of what is important in evaluating their teaching. Specific variables and loadings are provided in Table C-2.

Table C-2. Factors and Variable Loadings Related to Principals' Emphasis on Instruction or on Test Results

Items loading on factor	*Factor I – Emphasis on instruction*	*Factor II – Emphasis on test results*
Teacher content knowledge	.878	
Use of specific instructional strategies	.780	
Individualization of instruction	.773	
Student engagement	.750	
Instruction focused on raising test scores		.889
Student test results		.888

APPENDIX D. BAR CHARTS OF SIGNIFICANTLY DIFFERENT MEAN RATINGS ON DIFFERENT ASPECTS OF STANDARDS-BASED EDUCATION, BY SCHOOL CONDITIONS

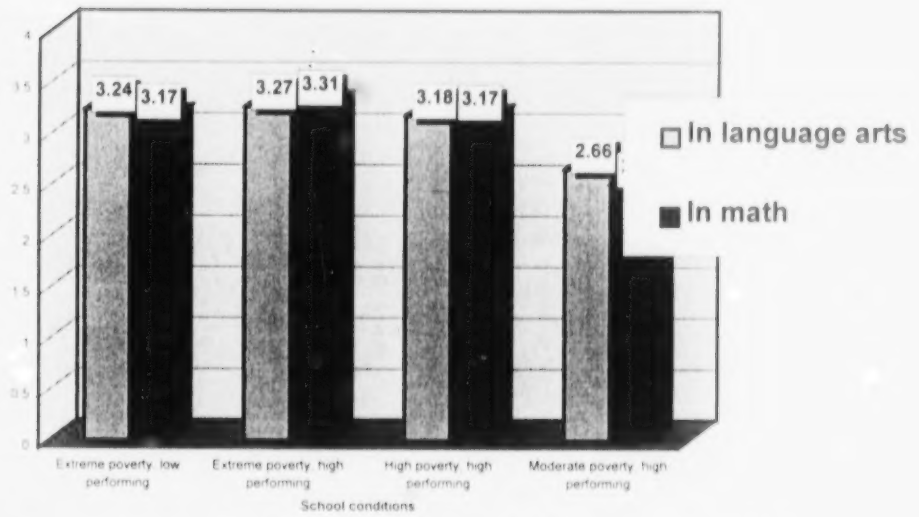

Chart D-1. Use of test results has improved learning, mean teacher ratings of agreement (1= strongly disagree, 4 = strongly agree) by school conditions

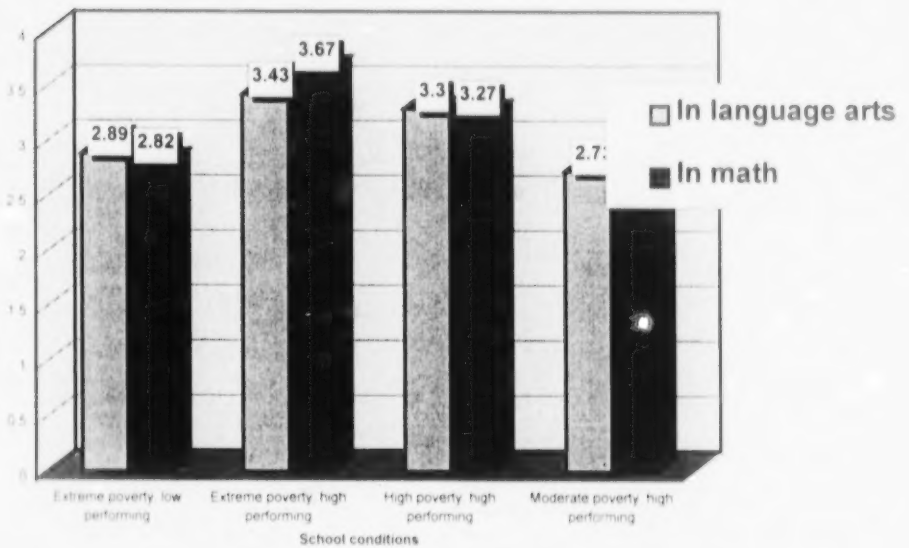

Chart D-2. Additional learning time is provided for nonproficient students, mean teacher ratings of agreement (1 = strongly disagree, 4 = strongly agree) by school conditions

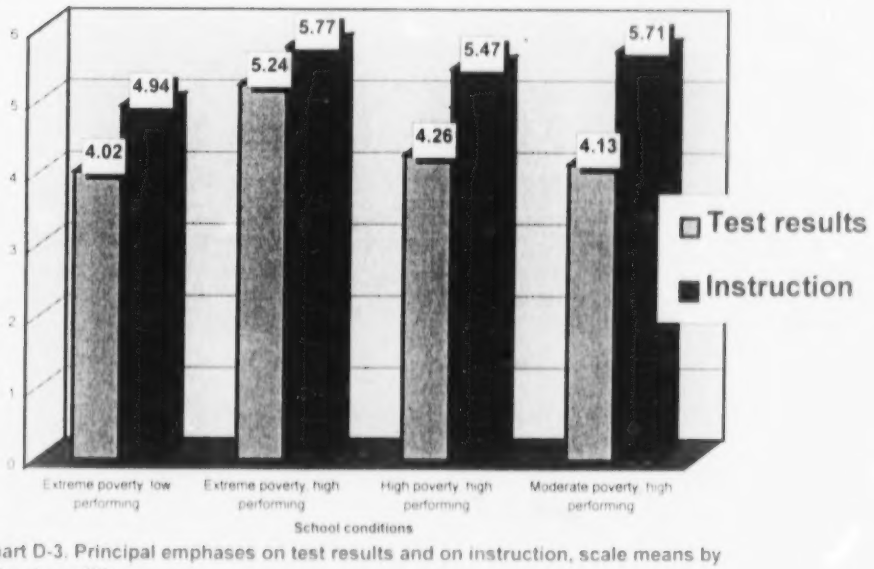

Chart D-3. Principal emphases on test results and on instruction, scale means by school conditions